LEARNING ASPECTS OF THE DIRECT STATISTICAL APPROACH TO THE OPTIMIZATION OF MONTE CARLO RADIATION TRANSPORT CALCULATIONS

K. W. BURN
ENEA (ERG–FISS)
BOLOGNA, ITALY

ABSTRACT

The Direct Statistical Approach (DSA) is a model for the optimization of parameters governing the control of the particle population in fixed source Monte Carlo particle transport calculations. It is currently applied to neutron and photon problems. The particle population is controlled at surfaces in phase space by means of the well known splitting and Russian roulette (RR) techniques with the amount of splitting/RR defined by the ratios of parameters ("importances") attached to the volumes ("cells") on either side of the surfaces. Biasing methods that change the weight of the particle tracks may be present but do not form part of the optimization, and splitting or RR is carried out independently of the weight of the track impinging on the surface. Thus if biasing is present, undesirable weight variations may still occur even with an optimum particle population control.

The DSA is based on a rigorous mathematical derivation of the second moment of the population and of the time per particle history. Under the above circumstances ("weight–independent splitting/RR") both these quantities are separable into a coefficient part that depends only on the characteristics of the particular problem (including possible biasing) and a function part that depends only on the phase space cell importances. It is thus a simple step to optimize the importances with the objective of minimizing the quality factor (the product of the second moment and time).

The coefficient parts of the second moment and time functions are estimated by Monte Carlo. Having established first estimates of these quantities, a minimization is carried out and a first set of "optimum" cell importances is found. These importances are then employed in a new Monte Carlo calculation that continues to accumulate estimates of the coefficients. At one point a second minimization is carried out and a second set of "optimum" cell importances is found that is then employed to further accumulate estimates of the second moment and time function coefficients, etc. This is the adaptive aspect of the DSA. Because of the statistical difficulty of estimating the second moment function coefficients (due to the necessity of estimating the square of the adjoint flux when a track crosses a surface), an approximation is employed. This approximation (named the "enhanced point–surface approximation") has proved to be excellent in most problems. Under these circumstances the above iterative learning procedure generally reaches the optimum in a few steps. The existence of a single integral parameter, the quality factor (that is estimated both directly and indirectly through the second moment and time function coefficients), allows the user to verify that the optimum set of cell importances has in fact been reached.

Recently with the objective of eliminating the weight variations present with biasing, a new DSA model has been developed that employs weight control rather than simple population control. In this model the amount of splitting/RR at a surface depends on the weight of the track impinging on the surface. In particular splitting or RR is carried out in such a way that the weight of the progeny that enter the new phase space cell is constant for that cell. These weights are the variables that it is desired to optimize. As is well known, under these circumstances the second moment is not separable. However an approximation (named the "non–integer approximation") has been introduced that renders the second moment separable. Under these circumstances the above phase space cell weights have been optimized employing a learning procedure for estimating the second moment and time function coefficients similar to the procedure for weight–independent splitting/RR. This new weight–dependent DSA model is currently being tested.

I. DELINEATING THE PROBLEM AND HISTORICAL DEVELOPMENT

The DSA ("Direct Statistical Approach") is a route towards optimizing (in a limited sense) Monte Carlo calculations and is currently employed on fixed source radiation transport problems. The radiation dealt with is at present neutrons and photons in the energy range < 20 MeV. The DSA employs as vehicle the code MCNP [currently version 4a – (Ref. 1)]. The DSA is due to Dubi, who in a series of original papers applied basic statistical considerations to the technique of surface splitting and Russian roulette (RR). This theoretical development culminated in a general model (Ref. 2).

As is well known splitting decreases the second moment whilst RR increases it. The other face of the coin is that splitting increases the CPU time whilst RR decreases it. There is an optimum amount of splitting: less and the saving in CPU time is more than offset by the rise in the second moment, more and the lowering in the second moment is more than offset by the rise in CPU time. The DSA describes in a precise mathematical fashion the second moment per source history S^2 (or the population second moment) and the time per source history τ, in terms of the splitting/RR parameters. The quality factor, q, defined as the product of S^2 and τ, is an appropriate measure of the inverse of the statistical efficiency of the Monte Carlo calculation (Ref. 3). Thus it is necessary to minimize q with respect to the splitting/RR parameters. This is possible because under the conditions described in Ref. 2 (the same amount of splitting or RR is carried out at a surface independent of the weight of the track arriving at the surface), both the second moment and time functions are, to all intents and purposes, separable into a problem–dependent coefficient part and a functional dependence on the splitting/RR parameters.

Over the last years the DSA has undergone further developments. These essentially concern substituting a closed–volume cell model for the original surface parameter model (Refs. 4, 5 and 6) and including all track bifurcations (as well as those due to the surface splittings) in the mathematical description of the second moment and time (Ref. 7). Also in Ref. 7 another limitation that previously did not allow tracks to experience splitting/RR at surfaces or cell boundaries after making a score, was removed. This seems to have marked a plateau in the development of the DSA weight–independent splitting/RR model. This model is summarized in the next paragraph.

Splitting/RR is executed at boundaries between closed volumes or cells in phase space (currently space and energy) according to the ratios of parameters ("importances") attached to the cells on either side of the boundary. Other variance reduction methods such as source or transport biasing may be present but are fixed (i.e. not optimized). The splitting/RR is independent of the weight of the track arriving at the boundary, thus it has purely the role of population control. The DSA aims to find the optimum set of cell importances. To do this the coefficients of the second moment and time functions are generated using Monte Carlo. A minimization code is then employed to find the set of cell importances that minimizes q (Ref. 6). This procedure is described later.

Of course it is not very efficient to split or Russian roulette tracks with widely varying weights by the same amount (at each cell boundary). When other variance reduction methods are present that alter the tracks' weight, a method such as the well known weight window is indicated that acts as a weight as well as a population control. To attempt to deal with problems involving weight variations other than those resulting from cell boundary splitting/RR, a new DSA model has been developed that splits and Russian roulettes at (phase space) cell boundaries according to the weight of the track arriving at the surface (Ref. 8). This model requires the weight after splitting or Russian roulette in each cell to be a particular value called the weight line or the starting weight. (This is different from the weight window which allows the weight to be in a particular range of values.) As before, other variance reduction methods that are present [biasing, DXTRAN spheres (Ref. 1), etc.] must have fixed parameters – the new DSA weight–dependent model aims to find the optimum set of cell starting weights only. Although under these circumstances the second moment is not separable, an approximation (named the "non–integer approximation") has been introduced that renders it separable and therefore makes the optimization practicable. This approximation represents a first approach. Notwithstanding, results so far look promising. Again the coefficients of the second moment and time functions are

generated using Monte Carlo and a minimization code is employed to find the set of cell starting weights that minimizes q.

II. THE SECOND MOMENT AND TIME FUNCTIONS

II.A. The DSA weight–independent cell importance model

The second moment, S^2, consists of three kinds of terms: Δ^o, Δ^i, and Γ representing the contributions from splittings at cell boundaries, volumetric bifurcations and direct detector scores respectively. (See Fig. 1 in which there is a general source and detector distribution that are not necessarily separated, and a general configuration of cells J or K with source cells j_0).

$$\Delta^o = \sum_{j_0} F(j_0) \sum_{I^i=1}^{L} \sum_{I^o=1}^{2N} \left\{ \text{coef–d}(j_0,I^i/I^o) \cdot \frac{1}{F(K)} . \frac{\text{int}(G(J,K))}{G(J,K)} . [2.G(J,K) - \text{int}(G(J,K)) - 1] \right\} \tag{1}$$

$$\Delta^i = \sum_{j_0} F(j_0) \sum_{I^i=1}^{L} \left\{ \text{coef–d}(j_0,I^i) \cdot \frac{1}{F(J)} \right\} \tag{2}$$

$$\Gamma = \sum_{j_0} F(j_0) \left\{ \text{coef–g}(j_0) \cdot \frac{1}{F(j_0)} + \sum_{I^o=1}^{2N} \left[\text{coef–g}(j_0,I^o) \cdot \frac{1}{F(K)} \right] \right\} \tag{3}$$

and

$$S^2 = \Delta^o + \Delta^i + \Gamma \tag{4}$$

Expressions (1), (2) and (3) are expressions (27), (30) and (36) respectively in Ref. 7.

In expression (1) each term within the summation brackets represents the contribution from splittings at the boundary between cells J and K with importances F(J) and F(K) respectively. [and G(J,K) = F(K)/F(J)]. The pair of variables: I^i/I^o represents the boundary between cells J and K, thus the double sum over I^i and I^o represents a sum over all cell boundaries.

In expression (2) each term within the summation brackets represents the contribution from bifurcations within each cell J. The variable I^i represents bifurcations of a certain type within cell J, thus the sum over I^i represents a sum over all bifurcation types and over all cells.

In expression (3) each term within the inner summation brackets represents the contribution from detector scores within each cell K. The sum over I^o represents a sum over all cells. The first term within the outer summation brackets represents the contribution from detector scores within each source cell j_0.

The coef–.... variables are problem–dependent coefficients whose values must be estimated:

– coef–d(j_0,I^i/I^o) in expression (1) involves the product of the square of the track's weight as it crosses a cell boundary and the square of the adjoint flux at the crossing point [see expressions (25) and (26) in Ref. 7].

– coef–d(j_0,I^i) in expression (2) involves the product of the square of the track's weight as it experiences a bifurcation within a cell and the product of the adjoint fluxes of each pair of tracks at the bifurcation point [see expressions (28), (29) and (23) in Ref. 7].

– coef–g(j_0) and coef–g(j_0,I^o) in expression (3) involve the square of a track's detector score [see expressions (35), and (32) and (33) in Ref. 7].

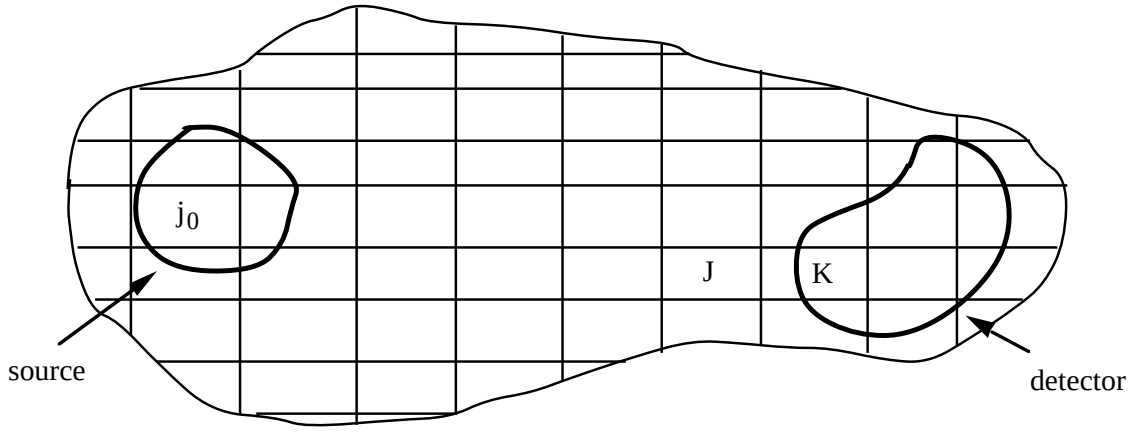

Fig. 1: General Cell Configuration

The estimates of these coefficients are all subject to greater statistical uncertainties compared with the estimate of a normal (linear) response as they include the squares of quantities. Estimating the square of the adjoint flux at a phase space point on a cell boundary, required in the coefficient: coef–$d(j_0,I^i/I^o)$, is particularly difficult. Therefore an approximation [the enhanced point surface (eps) approximation (see Ref. 8, Sec. IV)] has been introduced. This approximation has so far proved satisfactory in that the coefficients coef–$d(j_0,I^i/I^o)$ are still calculated accurately (at least in the vast majority of problems) but with much less statistical difficulty.

The time τ (or T in the nomenclature of Ref. 7) is given as follows:

$$T = \sum_{j_0} \frac{1}{F(j_0)} \left\{ \text{coef–t}(j_0) \cdot F(j_0) + \sum_{I^o=1}^{2N} \left[\text{coef–t}(j_0,I^o) \cdot F(K) \right] \right\} \tag{5}$$

Expression (5) is expression (43) of Ref. 7. Each term within the inner summation brackets represents the CPU time spent by tracks within each cell K. The sum over I^o represents a sum over all cells. The first term within the outer summation brackets represents the CPU time spent by tracks within each source cell j_0. The problem–dependent coefficients: coef–$t(j_0)$ and coef–$t(j_0,I^o)$ are much easier to estimate compared with the second moment coefficients as they consist only of linear or first moment–like quantities [see expressions (39), (40) and (42) in Ref. 7].

II.B. The DSA weight–dependent weight line model

After applying the non–integer approximation and in analogy with the previous Sec. II.A, the three kinds of terms composing the second moment, S^2, are given as follows:

$$\Delta_w^o = \sum_{j_0} \sum_{I^i=1}^{L} \sum_{I^o=1}^{2N} \left[\text{coef–d}_w(j_0,I^i/I^o)(1) \cdot \Omega(J) - \text{coef–d}_w(j_0,I^i/I^o)(2) \cdot \Omega(K) \right] \tag{6}$$

$$\Delta_w^i = \sum_{j_0} \sum_{I^i=1}^{L} \text{coef–d}_w(j_0,I^i) \cdot \Omega(J) \tag{7}$$

$$\Gamma_w \quad = \quad \sum_{j_0} \left\{ \text{coef}-g(j_0) \quad + \quad \sum_{I^o=1}^{2N} \left[\text{coef}-g_w(j_0, I^o) \; . \; \Omega(K) \right] \right\} \tag{8}$$

The time τ (or T_w in the nomenclature of Ref. 8) is given as follows:

$$T_w \quad = \quad \sum_{j_0} \left\{ \text{coef}-t(j_0) \quad + \quad \sum_{I^o=1}^{2N} \left[\text{coef}-t_w(j_0, I^o) \; . \; \frac{1}{\Omega(K)} \right] \right\} \tag{9}$$

Expression (6) combines expressions (27), (28), (31) and (33) of Ref. 8. Expression (7) combines expressions (36), (37) and (38) of Ref. 8. Expression (8) combines expressions (40), (41), (42) and (43) of Ref. 8. Expression (9) combines expressions (45), (46) and (47) of Ref. 8 with expression (42) of Ref. 7.

As far as the function dependence is concerned, comparing the weight–dependent expressions (6), (7), (8) and (9) with the weight–independent expressions (1), (2), (3) and (5), we see that the form is very different for the second moment contributions from the cell boundaries, Δ_w^o. [In particular we note the presence of negative terms in expression (6) which increase the statistical error on the estimates of this contribution.] Instead for the other second moment terms and the time, the form is similar once the cell starting weights $\Omega()$ are equated with the inverse of the cell importance F().

As far as the coefficients are concerned, the weight–dependent coefficients contain a different dependence on the surface–to–surface probabilities and accumulated weights compared with the weight–independent coefficients [see expressions (27), (28), (36), (40) and (45) in Ref. 8]. However the coefficients dealing with track bifurcations in expressions (6) and (7) still contain the square of the adjoint flux at the cell boundary crossing point and the product of the adjoint fluxes of each pair of tracks at the bifurcation point, respectively.

III. LEARNING THE SECOND MOMENT AND TIME FUNCTION COEFFICIENTS SO AS TO OPTIMIZE THE MONTE CARLO CALCULATION

The second moment and time function coefficients are estimated with Monte Carlo, using a modified form of MCNP–4a. They have not been written here for reasons of length. Some indications of their structure are given in Sec. II and they are found in Refs. 7 and 8 for the weight–independent and weight–dependent models respectively.

First moment–like quantities consist of multiplicative strings whose components are the product of the probability and the accumulated weight from (cell) boundary to boundary. The second moment and time function coefficients also consist of multiplicative strings and one component is always the boundary–to–boundary probability. Thus the normal probability density functions of the transport operator (biased as the case may be) can be used. Only the random variable (which for a first moment–like quantity is the track weight) is altered depending on which coefficient is estimated.

Estimating the coefficients with Monte Carlo has some general consequences:

1) The only information available is that gathered by the "estimating" tracks during their random walk (of course).

2) The well–known positive feed–back effects are present, especially marked when an iterative multi–step method is employed (see for example Ref. 9): that is, a chance low estimate of the importance (in the general sense) of some phase space region produces variance reduction parameters ("v.r." parameters) that militate against the sampling of that region in the next step. If as a consequence in the next step that region is not sampled, the v.r. parameters produced at the end of that step militate still more against the sampling of that region in the next step, etc.

3) The process must be started off by guessing the v.r. parameters to be used at the zero'th step.

As far as point 3) is concerned, the user must define a first set of v.r. parameters (cell importances or weight lines depending on which DSA model) to start the procedure off. What is a good set of v.r. parameters to estimate the second moment and time function coefficients (in particular the second moment) at the first and subsequent steps? Well we would intuitively expect that a good set of v.r. parameters would be one that tended to sample all phase space rather more, compared with the optimum set of v.r. parameters for the response of interest. That is, a set that forces the tracks towards the detector to a somewhat less extent compared with the optimum set of v.r. parameters. Notwithstanding, the optimum set of v.r. parameters generated at one step was employed at the next step. This has some advantages as will be shown.

The learning procedure will be illustrated through three sample problems. The first two are not very deep penetration whilst the third is. The first and the third employ the weight–independent model whilst the second employs the weight–dependent model.

The quantities of interest in the sample problems are as follows:

– the population first moment or response (R);

– the fractional standard deviation (fsd) on R;

– the population second moment (S^2);

– the total CPU time in minutes. [The CPU times refer to an IBM RISC/6000–990, with operating system: AIX/6000 ver. 3.2.5. The FORTRAN compiler was XL FORTRAN ver. 3 (including optimization).];

– the CPU time in minutes per source particle history (τ) (using NPS, the number of source particle histories);

– the quality factor ($q = S^2 \times \tau$);

– the inverse normalized quality factor ($= R^2/q$). {This quantity is used as a measure of the quality of the Monte Carlo calculation rather than the well–known the "figure of merit" (fom) [$= (CPU \times fsd^2)^{-1}$] (Ref. 1) because it does not include the book–keeping time for the accumulation of the second moment and time functions. Actually this time is not large so that the quantity R^2/q is quite close to the figure of merit.};

– the number of failures of the statistical checks on the response provided by MCNP4a (Ref. 1) ("stat fail" that may take the values 0 to 10). These useful checks provide an added insight into the statistical quality of the response and essentially concern the normality of the pdf of the mean. They are of course dependent on the CPU time and therefore not strictly a measure of the quality of the calculation in the sense of the previous paragraph.

The first sample problem is the well–known "Top Hat" (see Fig. 2). This is described in Refs. 8 and 10. No other variance reduction method is operating apart from cell boundary splitting/RR; therefore the DSA weight–independent model was employed. The results of the optimization procedure are shown in Table I (taken from Ref. 8).

The first remark to be made is that the second moment and time functions are accumulated from step to step, notwithstanding the fact that different v.r. parameters are employed at each step. (This is in contrast to the weight window generator in MCNP4a.) Thus the results in the columns in Table I entitled "Minimization of quality function" are all cumulative and (with the exception of R) are obtained from the second moment and time functions. The quantity R^2/q under the heading "init" has been evaluated at the current set of cell importances, whilst that under the heading "opt" has been evaluated at the optimum set of cell importances following operation of the minimization code. Instead the columns under the heading "Generation of second moment and time functions" are direct estimates obtained from the results of the particular run of the modified MCNP4a.

From Table I we see that there are a number of ways of verifying that the second moment and time function coefficients have converged to their proper estimates. Clearly at the optimum, operation of the minimization should not change much R^2/q. We see that at the zero'th step the functions indicate that it can be improved by slightly over a factor of 3 ($137.6 \rightarrow 430.4$), whilst at the fourth and final step the functions indicate that it can hardly be improved ($422.5 \rightarrow 423.2$). Furthermore the direct estimate of R^2/q at step n is to be compared with the function value R^2/q(init) at

step n, whilst the function value R^2/q(opt) at step n is to be compared with the direct estimate of R^2/q at step n+1. We see in Table I that the first comparison is not satisfied at steps 1 (compare 248.6 and 317.7) and 3 (compare 292.6 and 392.3) (and that this discrepancy is due to the second moment rather than the time) whilst the second comparison is not satisfied between steps 0 and 1 (compare 430.4 and 248.6) and between steps 2 and 3 (compare 429.4 and 292.6).

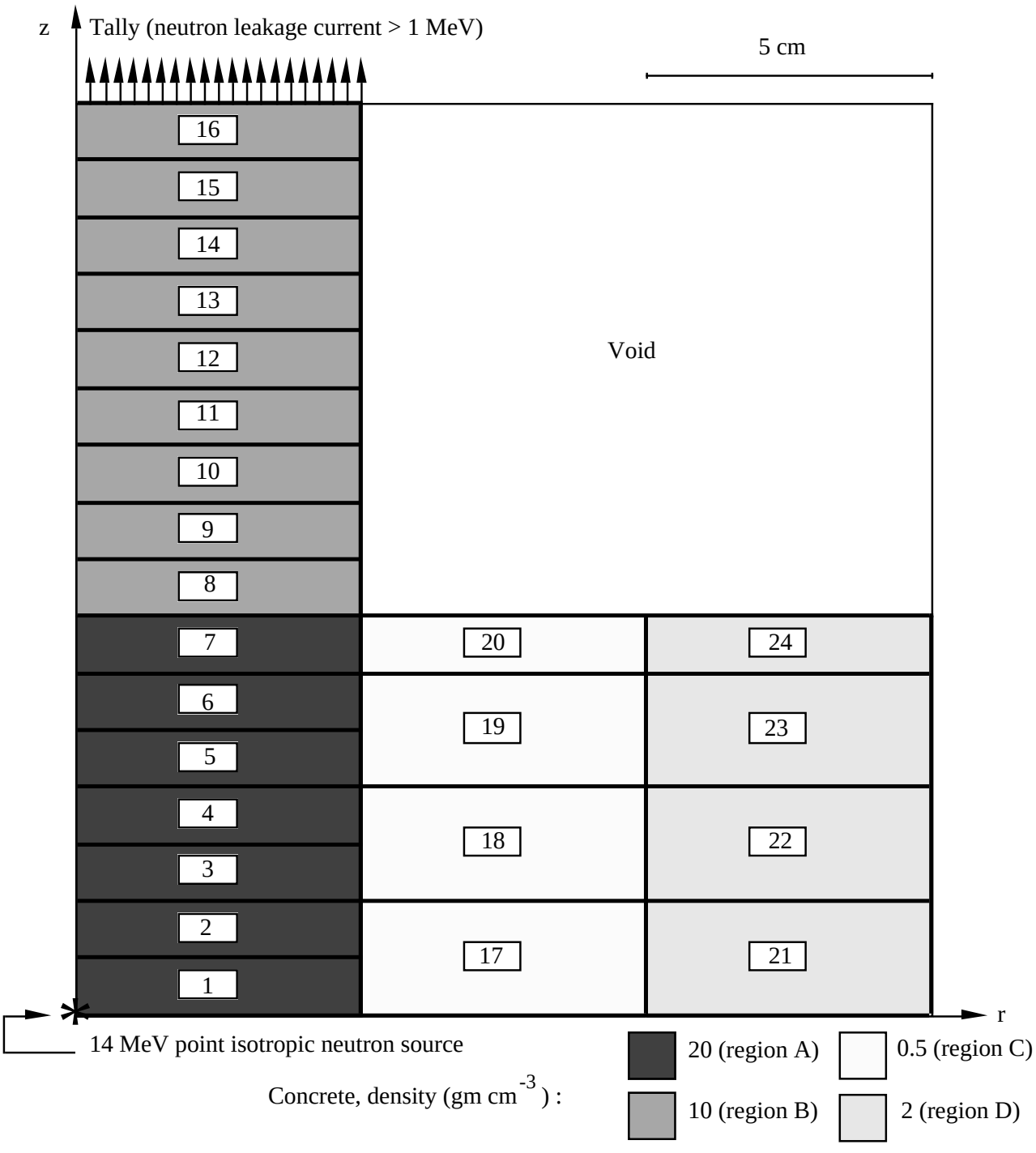

Fig. 2. Top Hat - geometry.

Thus the integral quantities: q and R^2/q allow us to know when we have generated reasonable second moment and time functions and therefore when we have arrived near the optimum. Examining the cell importances by themselves after each step (there are a total of 72 phase space cells in this problem as there are 3 energy groups) will not allow us to judge if we have arrived at the optimum or not. Clearly if the importance of some cell changes greatly from one step to the next, without the function information we are unable to know how important that change is to the quality of the calculation. In the absence of the functions the only information available is the direct estimate of the figure of merit,

equivalent to the direct estimate of R^2/q in Table I. (In fact the DSA functions contain a great deal more information – we are able if we wish to predict the second moment, time and calculational quality for any set of cell importances.)

Finally the fact that we have a good agreement between the direct estimate and the function value of R^2/q means that the eps approximation is good and the optimum reached is a real one. For this problem the DSA gives a calculational quality roughly double that of adjoint flux–based methods for controlling the population (see Ref. 10). It is thought that this is due to the two different contributing paths – through the high density regions A and B, and radially from region A to the low density regions C and D followed by streaming through the void towards the top of region B.

TABLE I

Top Hat. Iterative procedure to arrive at the optimum cell importances.

Step	Generation of second moment and time functions								Minimization of quality function					
	CPU (min)	R	fsd (%)	stat fail	S^2	τ	q	R^2/q	R	S^2	τ	q	R^2/q (init)	R^2/q (opt)
0	4.0	4.02–5	4.46	2	1.08–6	1.08–5	1.17–11	138.3	4.02–5	1.07–6	1.09–5	1.18–11	137.6	430.4
1	4.0	4.39–5	3.33	2	3.14–7	2.47–5	7.77–12	248.6	4.13–5	2.23–7	2.42–5	5.38–12	317.7	427.2
2	4.0	4.41–5	2.56	1	1.67–7	2.85–5	4.76–12	408.1	4.19–5	1.51–7	2.76–5	4.16–12	422.7	429.4
3	4.0	4.57–5	3.04	2	2.50–7	2.86–5	7.15–12	292.6	4.26–5	1.65–7	2.81–5	4.62–12	392.3	411.6
4	8.0	4.41–5	1.78	0	1.59–7	2.83–5	4.52–12	429.8	4.30–5	1.51–7	2.90–5	4.37–12	422.5	423.2

The second sample problem is a three leg penetration in high density concrete (see Fig. 3). This is described in Ref. 8. There is a DXTRAN sphere (Ref. 1) placed in the middle of each duct. These three spheres produce a wide range of track weights so a weight control is indicated, thus the DSA weight–dependent model was employed. The results of the optimization procedure are shown in Table II (taken from Ref. 8). The v.r. parameters are now the cell weight lines rather than the cell importances. Apart from this, the optimization procedure is similar to that of the DSA weight–independent model. In particular the optimum weight lines after step n are used in step n+1 to continue to accumulate the second moment and time function coefficients.

There are two seeming inconsistencies in the results in Table II:

– The function value R^2/q(init) (48.7) is not consistent with the function value of R (3.95E–8) and q (1.10E–16) at step 0. This is because at each step the functions are automatically adjusted before operation of the minimization code (by eliminating negative coefficients present in the second moment due to statistical error, making the second moment terms self–consistent and making the time function consistent with the second moment function). The function values of S^2, τ and q are given before adjustment whilst R^2/q(init) is given after adjustment. [In Ref. 8, in the equivalent Table XII, R^2/q(init) is given before adjustment.].

– The direct estimate and function value of the time at step 1 are rather different (1.42E–4 and 8.70E–5) although we expect the second moment rather than the time to exhibit possible differences. This discrepancy in the time is due to the fact that at the previous step 0, many cells gave no contribution to the second moment. Consequently these cells were eliminated from the optimization and were not assigned a weight line after step 0. At step 1 no splitting or Russian roulette game was played on tracks entering these cells. Although time is spent by tracks in these cells and is recorded in the direct estimate, this time is not recorded in the function value which is therefore an underestimate.

The main feature of Table II compared with Table I is an underestimation in the function value of the second moment compared with the direct estimate. [There are strong indications that this is due to the non–integer

approximation (see Ref. 8).] This gives rise to an overestimation in R^2/q. This has two consequences. Firstly we are unsure how far our optimum at step 5 is away from the real optimum. In Ref. 8 first indications (on the Top Hat problem) are that we are not far from the real optimum. Secondly we are unable to control the convergence of the optimization procedure in as close a way as in the previous problem as we cannot compare the direct estimate and function value of R^2/q as we could before. However we do know that the function value of R^2/q is always an overestimate because the second moment is always underestimated and experience so far is that function value of the second moment has never been less than 60% of the direct estimate.

For this problem the weight–dependent DSA gives a calculational quality similar to that of adjoint flux–based methods for controlling the track weight (see Ref. 8).

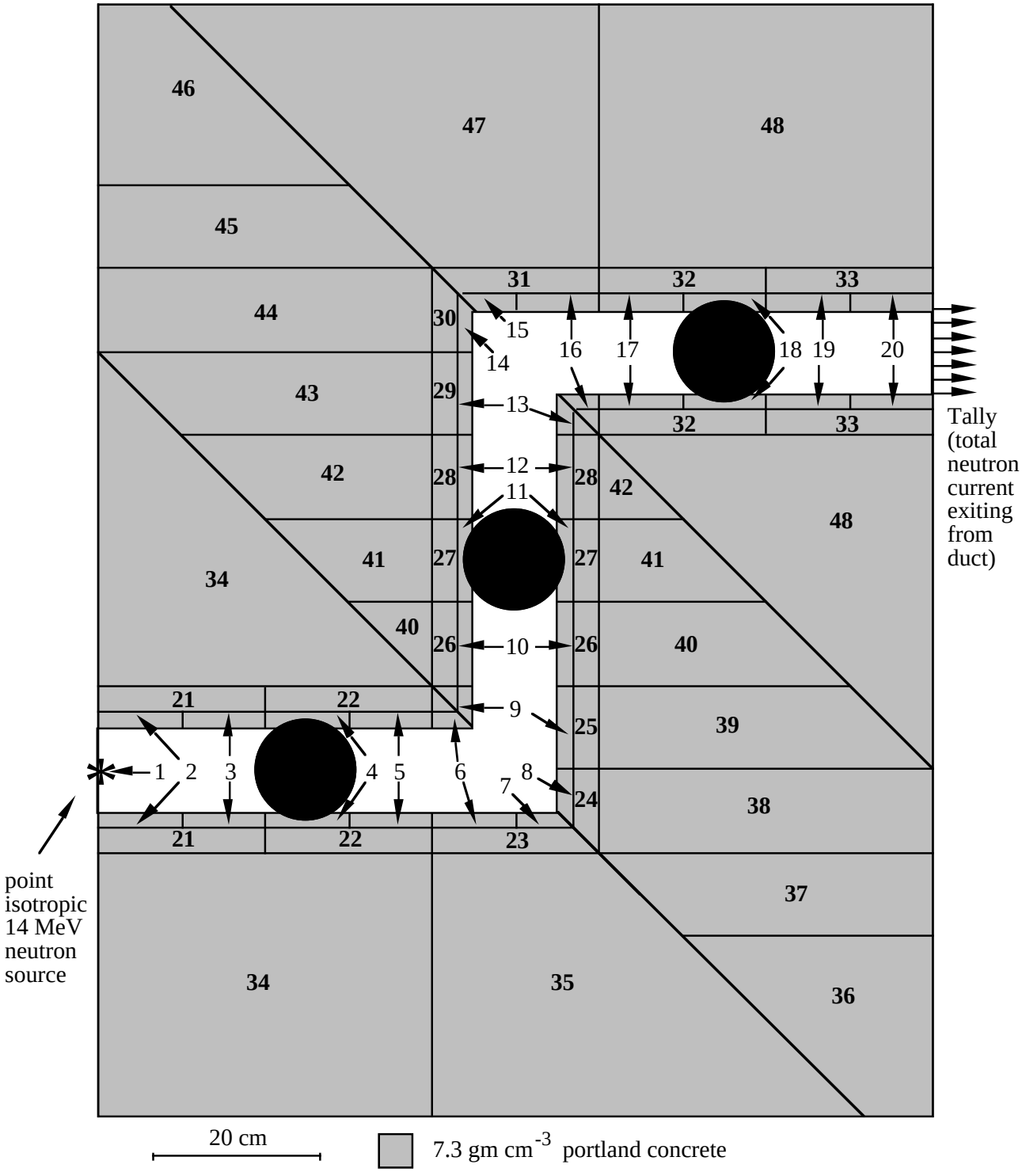

Fig. 3. Dog leg penetration in high density concrete – geometry.

TABLE II

Dog leg penetration in high density concrete. Iterative procedure to arrive at the optimum starting weights.

Step	Generation of second moment and time functions								Minimization of quality function					
	CPU (min)	R	fsd (%)	stat fail	S^2	τ	q	R^2/q	R	S^2	τ	q	R^2/q (init)	R^2/q (opt)
0	6.0	3.95–8	12.1	2	5.14–13	2.42–4	1.25–16	12.5	3.95–8	4.54–13	2.42–4	1.10–16	48.7	371.6
1	4.0	3.22–8	8.39	1	1.69–13	1.42–4	2.40–17	43.1	3.58–8	9.63–14	8.70–5	8.38–18	154.4	226.0
2	7.0	3.05–8	3.66	0	6.23–14	1.22–4	7.60–18	122.6	3.31–8	5.14–14	1.19–4	6.09–18	180.0	187.0
3	4.0	3.30–8	4.73	0	6.75–14	1.31–4	8.85–18	122.8	3.30–8	4.55–14	1.35–4	6.17–18	177.0	181.9
4	10.0	3.36–8	3.27	4	7.92–14	1.40–4	1.11–17	101.8	3.32–8	5.27–14	1.37–4	7.22–18	153.2	174.9
5	10.0	3.29–8	2.87	0	6.20–14	1.35–4	8.38–18	129.5	3.32–8	4.83–14	1.34–4	6.47–18	169.9	170.1

With the relatively high figures of merit (about 400 and 100) in the first two sample problems, it was easier to demonstrate the operation of the optimization procedure. In contrast the third sample problem is a deep penetration calculation. It is a simplification, whilst retaining the variance characteristics, of a real problem of calculating the activation rate in a wire in the core of a WWRSM–type research reactor. The geometry is shown in Fig. 4. (The wire in Fig. 4 is not to scale – it has a diameter of 0.5 mm.) The outer boundaries are reflection surfaces. The fission source is fixed and is distributed homogeneously outside the detector channel. Neutron production from fissions is suppressed.

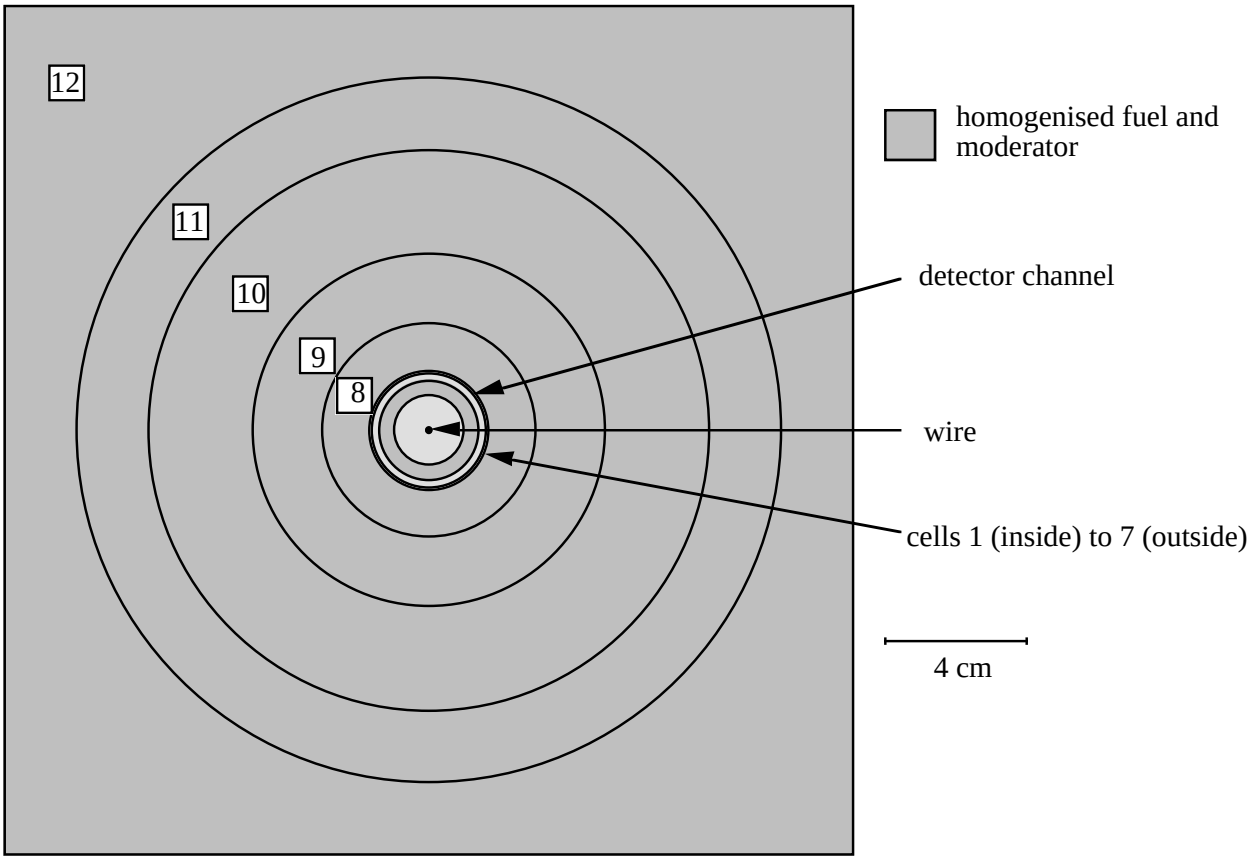

Fig. 4. Activation rate in wire in reactor core – geometry

The thinness of the wire is the first factor that contributes to the penetration. The response of interest is the sum of the (n,γ) reaction rates in ^{94}Zr and ^{96}Zr. The second factor that contributes to the penetration is the resonance structure of both the ^{94}Zr(n,γ) and the ^{96}Zr(n,γ) cross–sections. The ^{96}Zr(n,γ) cross–section (which accounts for approximately 90% of the total response) is shown in Fig. 5. A total of 15 energy groups were employed with a structure that took into account both the resonances in Fig. 5 as well as those in the ^{94}Zr(n,γ) cross–section. The upper limits of these groups were as follows: 20 MeV, 14.5 keV, 13 keV, 7.8 keV, 7.5 keV, 5.7 keV, 5 keV, 4.9 keV, 4.2 keV, 3.7 keV, 2.5 keV, 2 keV, 325 eV, 280 eV and 10 eV.

As there is no weight variation in this problem apart from that due to splitting/RR at the cell boundaries, the DSA weight–independent model was used. The optimization procedure is shown in Table III. Studying Table III we see that:

1) It is very difficult to arrive at the optimum and even after step 11 (with a total of 106 hours of accumulated CPU time) the procedure has still not converged.

2) The direct estimate of R^2/q (under "Generation of second moment and time functions") does not at first sight seem to be improving as the optimization progresses. However when we examine the response value R under the same heading, we see that in the first steps, relatively high values of R^2/q correspond to underestimates in R. Only in the last three steps do we obtain both a relatively high value of R^2/q and a R that does not seem an underestimate.

Fig. 5: The 96Zr(n,gamma) cross-section

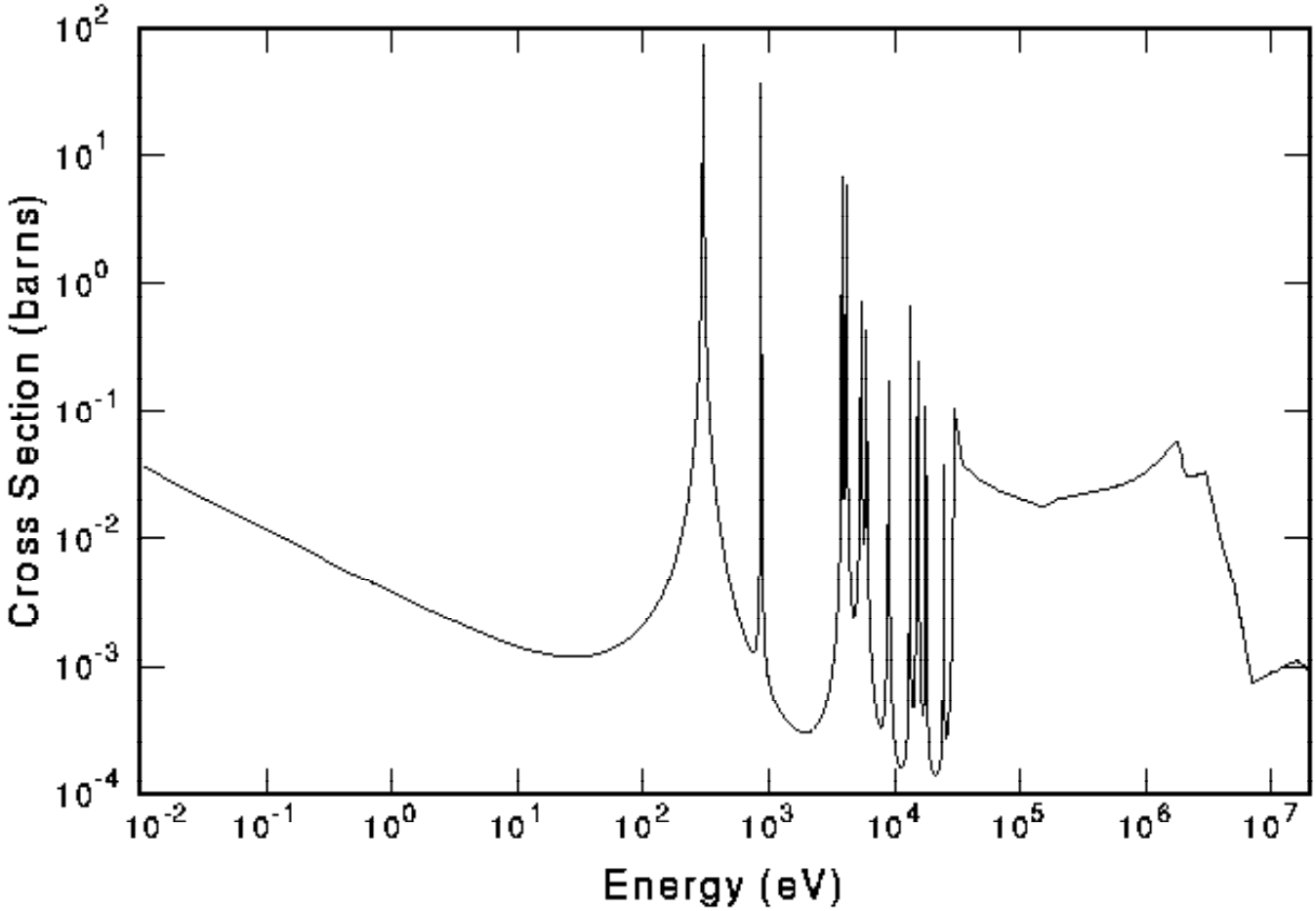

TABLE III

Activation rate in wire in reactor core. Iterative procedure to arrive at the optimum cell importances.

Step	Generation of second moment and time functions								Minimization of quality function					
	CPU (min)	R	fsd (%)	stat fail	S^2	τ	q	R^2/q	R	S^2	τ	q	R^2/q (init)	R^2/q (opt)
0	24.0	9.96–6	22.35	3	8.41–7	1.20–4	1.01–10	0.98	9.96–6	8.41–7	1.20–4	1.01–10	0.98	253.9
1	20.1	9.74–5	65.31	7	4.70–4	1.41–4	6.64–8	0.14	4.55–5	1.95–4	1.38–4	2.69–8	0.077	52.0
2	38.0	2.99–5	16.45	2	1.46–6	5.14–4	7.52–10	1.19	4.28–5	2.15–7	5.57–4	1.20–10	15.3	85.0
3	32.0	2.56–5	13.21	3	1.06–6	2.83–4	3.00–10	2.19	3.92–5	1.27–7	3.00–4	3.82–11	40.1	58.4
4	36.1	2.77–5	24.57	8	3.44–6	3.97–4	1.37–9	0.56	3.75–5	4.96–7	4.17–4	2.07–10	6.8	42.0
5	212.1	3.47–5	7.37	2	2.27–6	5.05–4	1.15–9	1.05	3.64–5	2.14–7	5.13–4	1.10–10	12.1	30.0
6	76.1	3.69–5	9.47	2	1.39–6	5.49–4	7.62–10	1.79	3.65–5	1.34–7	5.54–4	7.44–11	17.9	26.0
7	100.1	3.24–5	12.22	4	2.63–6	4.93–4	1.29–9	0.81	3.59–5	2.21–7	4.96–4	1.09–10	11.8	24.7
8	540.0	4.44–5	10.79	8	2.05–5	4.97–4	1.02–8	0.19	3.96–5	8.16–6	4.96–4	4.05–9	0.39	19.2
9	1079.	3.78–5	3.20	1	2.71–6	4.79–4	1.30–9	1.10	3.88–5	2.52–7	4.73–4	1.19–10	12.6	18.4
10	1024.	3.88–5	4.54	6	5.35–6	4.88–4	2.61–9	0.58	3.88–5	6.72–7	4.68–4	3.15–10	4.77	15.9
11	3181.	3.80–5	1.93	2	2.79–6	5.05–4	1.41–9	1.03	3.84–5	4.94–7	5.02–4	2.48–10	5.95	15.4

3) At each step there is an indicated improvement in the functional value of R^2/q that does not seem to be borne out with the direct estimate. (This is another sign that the optimization procedure has not converged yet, or in other words that the estimate of the second moment coefficients is still poor.)

4) Notwithstanding the fact that the weight–independent DSA model is employed so that no approximation is necessary to render the second moment separable, there is a very substantial difference between the direct estimate and the function value of the second moment. (At step 11 the function value is a factor 5.6 less than the direct estimate.) This discrepancy can only be due to the eps approximation. It is suspected that the resonance structure in both the ^{94}Zr(n,γ) and ^{96}Zr(n,γ) cross–sections is of such a complexity that the binning of the energy into 15 groups is not sufficient and is responsible for the breakdown in the approximation.

Concerning point 4), the fact that the function value of the second moment is so badly underestimated, leads us to wonder whether the "optimum" is also poor. It is interesting to compare this optimum with that of approaches based on the inverse of the adjoint flux estimate. To this end, weight windows were generated (see Ref. 1) at three different points during the optimization procedure: at steps 8, 9 and 11. The CPU times of these steps: 540, 1079 and 3181 min. respectively (see Table III) was the same as the total cumulative CPU time of the optimization procedure up to but not including the step in question. This just means that an equal amount of CPU time was spent on generating the DSA function coefficients as was spent on generating the weight window. (If anything the quality of the CPU time spent on generating the weight window was better than that spent on generating the DSA information if we assume that parameters that are good for the first moment are also good for the second.) The comparison between the DSA runs and the runs with the generated weight window ("gen. ww") is shown in Table IV. (In Table IV the fom rather than R^2/q is used as a measure of the quality of the calculation.) We see that there is a difference in quality of a factor of 2, 9 and 8 at

the steps 8, 9 and 11 respectively. However at step 8 the result with the generated weight window looks an underestimate so the factor 2 is unreliable. Thus notwithstanding the dire underestimate in the second moment function value and the fact that the optimization procedure does not yet seem to have converged, the DSA is performing relatively well with respect to an adjoint flux–based method.

TABLE IV

Activation rate in wire in reactor core. Comparison of DSA results with weight window generator results.

step	CPU (min)	NPS	R	fsd (%)	stat fail	fom
8 (DSA)	540.0	893398	4.44–5	10.79	8	0.16
(gen. ww)	1078.6	370261024	3.24–5	10.34	4	0.087
9 (DSA)	1078.5	1843940	3.78–5	3.20	1	0.90
(gen. ww)	1080.0	300535787	3.57–5	9.72	3	0.098
11 (DSA)	3181.1	5160311	3.80–5	1.93	2	0.84
(gen. ww)	4875.2	1641270050	3.95–5	4.25	0	0.11

To try to understand the reason behind the difference in the quality of the results from the two methods, in Table V are shown the DSA cell importances ("DSA ci") employed at step 11 for a sample of three energy groups (the lower two containing important resonances) in each of the 12 spatial cells (see Fig. 4). The lower bounds of the generated weight window values were inverted and normalized to the DSA cell importance in the energy group 325–280 eV in cell 8. Comparing the resulting values ("1/wwlb") with the DSA cell importances we see that the weight window generator produces a much higher multiplication in the inner part of the detector channel (cells 1 to 4) compared with the DSA. Such effects have also been noted with streaming problems in ducts (see Ref. 10).

TABLE V

Activation rate in wire in reactor core. Comparison of DSA cell importances at step 11 with generated weight window.

cell	> 14.5 keV		5.0 – 4.9 keV		325 – 280 eV	
	DSA ci	1 / wwlb	DSA ci	1 / wwlb	DSA ci	1 / wwlb
1	6.95+1	1.57+2	4.24+3	1.99+4	2.76+4	2.78+5
2	1.39+1	1.89+1	1.21+2	2.91+2	1.38+4	1.53+4
3	6.95+0	1.37+1	1.21+2	2.74+2	1.28+4	7.64+3
4	2.02+1	1.52+1	1.21+2	6.99+1	1.28+4	2.29+3
5	3.11+0	1.48+1	1.21+2	1.78+2	6.39+3	2.31+3
6	1.56+0	1.45+1	3.34+1	1.17+2	6.39+2	1.76+3
7	1.56+0	1.46+1	6.35+0	1.25+2	3.20+2	1.69+3
8	7.79–1	1.26+1	6.35+0	6.53+1	3.20+2	3.20+2
9	7.79–1	9.94+0	3.11+0	5.10+1	3.20+2	4.25+1
10	7.79–1	5.84+0	7.79–1	1.02+1	1.86+1	1.49+0
11	3.89–1	4.17+0	3.34–1	2.04+0	1.86+1	4.75–1
12	3.89–1	3.22+0	1.13–1	7.58–1	1.74–2	1.77–1

IV. POINTS FOR DISCUSSION

1) Might it be the case that the relatively complex form of the dependence of the quality function on the variance reduction parameters [especially in the case of the weight–independent splitting/RR model – see expressions (1) to (5)], makes the resulting "optimum" variance reduction parameters less sensitive to statistical errors in the information (the second moment and time function coefficients) that the optimizing method requires? Could this effect offset the statistical difficulty of estimating the second moment function coefficients?

2) The DSA models are at present written for fixed source problems only. The third sample problem in Sec. III ("Activation rate in wire in reactor core") was originally a k_{eff} calculation. However, referring to the source iteration method for estimating k_{eff}, the DSA, as it is currently implemented, treats each fission source generation independently instead of passing the information from one generation through the fission source to the next generation. Under the latter circumstances, over a number of generations, correlations will develop: neutrons giving detector contributions in one generation will undergo fission and the resulting fission neutrons will give detector contributions, etc. As the DSA is in principle able to take such correlations into account, it would be interesting to compare it with adjoint flux–based approaches on such problems.

3) It is intended in the future to study the operation of variance reduction in high energy transport problems (GeV) in which there is a large amount of natural bifurcation (cascades or showers). This bifurcation contributes to the second moment and is explicitly taken into account by the DSA. Experience so far in problems with substantial volumetric bifurcation in the energy range < 20 MeV [usually due to the presence of DXTRAN spheres (Ref. 1)] is that this is not the major contribution to the second moment.

4) The solid mathematical basis of the DSA has the consequence of allowing its application to fields outside the one for which it was specifically developed. In particular two possible fields of application have appeared:

– The problem of optimizing variance reduction parameters to more than one response, with the objective of achieving a comparable (and minimum) fractional error associated with each response in a given CPU time. A possible approach is to apply the DSA after forming a compound second moment function by summing the second moment functions of each individual response, having normalized each function by dividing the coefficients by the squares of the respective first moments.

– For some problems (e.g. cosmic ray showers in the earth's atmosphere), the second (and higher) moments of a response associated with the shower are themselves quantities of interest as well as the first moment. Currently moments higher than the first can only be estimated by means of an analogue calculation, which can be extremely long. With the DSA, the second moment may be found in a calculation employing either weight–independent or weight–dependent splitting/RR between space/energy cells. Instead as far as moments higher than the second are concerned, the theory needs to be extended before they can be estimated in an analogous fashion.

V. SUMMARY

Operation of the Direct Statistical Approach (both the weight–independent and weight–dependent models) with particular reference to its learning aspects, has been demonstrated through three sample problems. Some comparisons with methods for population and weight control based on the inverse of the adjoint flux estimate have been made.

ACKNOWLEDGEMENT

Many thanks to Dr. Gabor Hordòsy (KFKI–AEKI, Budapest) for supplying the specifications of the original model for the third sample problem (activation rate in wire in the core of a WWRSM–type reactor).

REFERENCES

1. “MCNP – A General Monte Carlo N–Particle Transport Code,” LA–12625–M, J.F. BRIESMEISTER, Ed., Group XTM, Los Alamos National Laboratory (1993).

2. A. DUBI, "General Statistical Model for Geometrical Splitting in Monte Carlo - parts I and II," *Trans. Theory and Stat. Phys.* **14–2** 167 and 195 (1985).

3. A. DUBI, T. ELPRIN and D. J. DUDZIAK "Geometrical Splitting in Monte Carlo," *Nucl. Sci. Eng.*, **80**, 139 (1982).

4. K. W. BURN, “Extension of the Direct Statistical Approach to a Volume Parameter Model (Non-Integer Splitting),”, *Ann. nucl. Energy* **17-6** 293 (1990).

5. K. W. BURN, "Optimizing Cell Importances using an Extension of the DSA – Theory, Implementation, Preliminary Results," *Progress in Nuclear Energy* **24**, 39 (1990).

6. K. W. BURN, "Complete Optimization of Space/Energy Cell Importances with the DSA Cell Importance Model," *Ann. nucl. Energy* **19–2** 65 (1992).

7. K. W. BURN, "Extending the Direct Statistical Approach to include Particle Bifurcation between the Splitting Surfaces," *Nucl. Sci. Eng.*, **119**, 44 (1995).

8. K. W. BURN, "A New Weight–Dependent DSA Model," submitted for publication to *Nucl. Sci. Eng.*, (January, 1996).

9. T. E. BOOTH, “Zero–Variance Solutions for Linear Monte Carlo," *Nucl. Sci. Eng.*, **102**, 332 (1989).

10. T. E. BOOTH and K. W. BURN, "Some Sample Problem Comparisons Between the DSA Cell Model and the Quasi–Deterministic Method," *Ann. nucl. Energy* **20–11** 733 (1993)